Introdução à Cointegração: abordagem uni-equacional

Artur Silva Lopes
ISEG, ULisboa

26 de Abril de 2020, versão 1.02

Conteúdo

Prefácio

A maior parte deste texto já tem cerca de 20 anos. Ele resulta de apontamentos de aulas que foram escritos por volta do ano 2000 e que tiveram poucas alterações ao longo dos anos, até porque ao invés de, como se previa, se expandir, o tempo lectivo que foi possível afectar a esta matéria se manteve (ou até se reduziu). Por outro lado, só agora houve disponibilidade de tempo para passar para o computador estes apontamentos, que se destinam a apoiar a leccionação da disciplina de Macroeconometria I. A não inclusão da abordagem multi-equacional, de sistema, deriva do "tratado de Tordesilhas" que foi estabelecido no início deste século com o colega da altura de Macroeconometria II.

É devido à passagem de tanto tempo sobre as primeiras versões que é frequente não poder indicar a referência concreta em que me baseei para este ou aquele tópico. Penso, no entanto, que todos os textos mais importantes que empreguei são indicados na bibliografia.

A matéria da cointegração representa o coroar da grande revolução (das raízes unitárias e da cointegração) que foi iniciada, muito tenuemente, ainda no final dos anos 70, e que terá atingido o auge cerca do ano 2000. Afectou profundamente a macroeconometria e é muito difícil descrever correcta, justamente e em poucas palavras as formas como isso ocorreu. Um indicador significativo é o facto de Clive W. J. Granger ter recebido o prémio Nobel da economia em 2003 pelo seu trabalho inovador e revolucionário neste domínio. E em 2003 já era mais que merecido.

Se a matéria das raízes unitárias representa a perspectiva univariada da modelação da generalidade das séries macroeconómicas, a da cointegração representa o seu culminar, agora sob o ponto de vista multivariado (e, por esse facto, com mais interesse para a análise económica). Talvez não muito estranhamente, contudo, também a aplicação da abordagem e das técnicas da cointegração vem conhecendo um retrocesso significativo na última dúzia de anos. O longo prazo parece ter deixado de ser importante.

Este texto representa também o culminar do reduzido programa da disciplina de Macroeconometria I. O seu carácter é, por isso, mais literário, menos algébrico ou aparentemente menos técnico que o usual. É o estuário em que as várias correntes dos capítulos anteriores vêm desaguar. É o encaixar final de várias peças, com a restrição já referida, e sem poder contar com a importante peça da não-linearidade na média. Como, por outro lado, se trata de uma primeira versão, agradecem-se as indicações de erros, falhas e omissões. Sugestões também são bem-vindas.

25 de Abril de 2019

1 Introdução

Nesta introdução começa por se apresentar um conjunto de pequenas propriedades úteis e uma primeira definição de cointegração, que também se considera útil para introduzir o tema.

1.1 Propriedades básicas

Enuncia-se em seguida um conjunto de pequenas propriedades básicas, bastante triviais, quase axiomas, mas bastante úteis para fazer alguns raciocínios[1]. Considerem-se duas constantes não nulas, a e b. Então:

i. se $x_t \sim I(0)$, tem-se que $(a + b\,x_t) \sim I(0)$;
   se $x_t \sim I(1)$, tem-se que $(a + b\,x_t) \sim I(1)$;

ii. se x_t e y_t são ambas $I(0)$, tem-se que $(a\,x_t + b\,y_t) \sim I(0)$, ou seja, qualquer combinação linear de processos $I(0)$ ainda é $I(0)$;

iii. se $x_t \sim I(1)$ mas $y_t \sim I(0)$, tem-se que $(a\,x_t + b\,y_t) \sim I(1)$, ou seja, a propriedade $I(1)$ é dominante em relação à $I(0)$;

iv. *em geral*, se $x_t \sim I(1)$ e $y_t \sim I(1)$, tem-se que $(a\,x_t + b\,y_t) \sim I(1)$, ou seja, *em geral* (mas não sempre), qualquer combinação linear de séries $I(1)$ ainda é $I(1)$.

Insiste-se que a última não é, efectivamente, uma propriedade, pois não é universal; é apenas uma constatação de algo que é muito frequente mas que nem sempre se verifica.

1.2 Cointegração: primeira definição

Introduz-se já uma primeira definição, limitada e preliminar, de cointegração, apenas para o caso de duas séries integradas de primeira ordem, $I(1)$.

Definição: se $x_t \sim I(1)$ e $y_t \sim I(1)$, mas existe uma combinação linear, $(y_t - \beta x_t) = u_t \sim I(0)$, então x_t e y_t dizem-se cointegradas, $(x_t, y_t) \sim CI(1,1)$, e o vector $[1 \quad -\beta]'$ é chamado de vector de cointegração.

[1] Estas propriedades aparecem, por exemplo, na introdução do livro de Engle e Granger (1991).

Note-se que só se exige que a combinação linear seja I(0); ela pode ter média não nula ou pode até ter uma tendência, i. e., ser estacionária em tendência. Todavia, se para além de ser $I(0)$, se pretender que u_t tenha ainda média nula (portanto, sem tendência), pode ser necessário introduzir uma constante e/ou um termo de tendência linear na relação, isto é, pode ser necessário considerar

$$y_t - \gamma_0 - \gamma_1 t - \beta_t x_t = u_t \sim I(0),$$

para que u_t tenha média nula.

Como é então evidente, uma relação de cointegração é especial por ser pouco frequente, por ser uma excepção à regra geral enunciada acima. Note-se ainda que, se y_t e x_t são ambas $I(1)$, então, pela decomposição de Beveridge e Nelson, ambas contêm uma tendência estocástica. Ora, se $(y_t, x_t) \sim CI(1,1)$, isto é, se $y_t - \beta x_t = u_t \sim I(0)$, é porque a tendência estocástica desapareceu com a combinação linear, foi aniquilada; então, é porque ela era comum às duas séries. Isso significa então que, embora ambas aparentem evoluir de forma arbitrária, sem qualquer atracção por qualquer valor ou sem um padrão sistemático de comportamento, parecendo vaguear sem destino, o seu comportamento de longo prazo é conduzido ou guiado pela mesma tendência estocástica; ou seja, no longo prazo, elas covariam conjuntamente, de forma a satisfazer essa relação (linear). Elas evoluem tendendo a acompanhar-se mutuamente.

Em termos humorísticos, não pode deixar de se referir a excelente analogia entre a evolução de 2 séries cointegradas e os trajectos de um bêbado e do seu cão, saindo de um bar e regressando a casa durante a madrugada, inventada por Murray (1994). Recomenda-se a leitura do pequeno artigo mas, mais ainda, o recurso à imaginação.

No gráfico apresentado mais adiante, na subsecção 5.1, apresentam-se exemplos simulados de duas séries I(1) e não cointegradas, no painel superior esquerdo, e cointegradas, no painel superior direito.

2 Cointegração e teoria económica

A teoria económica sugere frequentemente a existência de relações entre variáveis, que podem ser interpretadas como relações de equilíbrio de longo prazo. Isto significa que, em termos empíricos, a observação do seu comportamento em períodos de tempo curtos é susceptível de revelar apenas desequilíbrios, situações em que a relação não se verifica. Todavia, no longo prazo, sobre horizonte temporais longos, de bastantes anos, as variáveis tenderão a mover-se de forma a satisfazer, de forma aproximada, não necessariamente exacta, a referida relação, com desvios que são estacionários, que não tendem a divergir, que não tendem a acumular-se de forma sistemática. É como se, em situações de grandes desequilíbrios ou após alguns períodos de desequilíbrio, existissem forças ou mecanismos que actuam no sentido de restaurar o equilíbrio, de o reestabelecer. Assim, uma relação de equilíbrio de longo prazo *tende* a verificar-se de forma *aproximada*, *em média*, *no longo prazo*.

Note-se que esta é uma perspectiva ou visão de equilíbrio estatística; não é a do sentido económico usual, de se entender que a situação de equilíbrio é aquela em que existe igualdade entre quantidades desejadas e quantidades transaccionadas. Ou seja, o sentido estatístico ou econométrico é de uma relação sobre um horizonte temporal longo, de erros de equilíbrio ou desequilíbrios meramente temporários, isto é, sem tendência para se acumularem, e variando sempre dentro de certos limites.

A cointegração traduzirá assim uma relação de equilíbrio de longo prazo, de co-movimento sistemático de longo prazo entre as variáveis envolvidas.

Formalmente, poder-se-ia considerar que, com $y_t \sim I(1)$ e $x_t \sim I(1)$, o equilíbrio exigiria que $y_t = \beta\, x_t, \forall t$, isto é, que teria que se ter $y_t - \beta\, x_t = u_t \equiv 0, \forall t$. Mas esta é uma condição extremamente exigente, que nunca seria observada empiricamente.

De outra forma, se, apesar de individualmente as duas séries serem $I(1)$, se tem que $y_t - \beta\, x_t = u_t \sim I(0)$, preferencialmente com média nula, isso significa que u_t pode ser interpretado como o erro ou desvio de equilíbrio, como desequilíbrio, pois:

a) o comportamento de reversão para a média (*mean reversion* ou *regression*) assegura que ele volta frequentemente a essa média, nula, isto é, que a situação de equilíbrio actua como um atractor para as duas séries;

b) que é meramente temporário ou transitório, sem apresentar tendência para divergir ao longo do tempo;

c) como tem variância finita e constante, mantém-se sempre dentro de certos limites, isto é, flutua sempre dentro de uma banda de largura constante no tempo.

Vejamos agora alguns exemplos simples de hipóteses ou teorias económicas que se traduzem estatisticamente em relações de cointegração.

Exemplo 1: hipótese do rendimento permanente do consumo. Uma das implicações mais conhecidas desta hipótese ou teoria, formulada por M. Friedman, é de que as variações do consumo das famílias em cada período são imprevisíveis. Recorde-se que, segunda ela, o consumo dos indivíduos em cada período é determinado não apenas pelo seu rendimento no período mas pela sua expectativa de longo prazo de rendimento, o chamado rendimento permanente.

Ora, uma versão simples da hipótese implica que o consumo privado ou das famílias tem dois componentes,

$$c_t = c_t^P + c_t^T,$$

onde a primeira componente é a permanente ou de longo prazo e a segunda é a transitória. Assumindo que $c_t^T \sim I(0)$ e como, segundo a hipótese, $c_t^P = \beta\, y_t^P$, onde y_t^P representa o referido rendimento permanente, tem-se

$$c_t = \beta\, y_t^P + c_t^T,$$

que, desde que $y_t^P \sim I(1)$,[2] representa uma equação de cointegração, com c_t^T a representar o erro de equilíbrio estacionário.

Exemplo 2: hipótese de eficiência dos mercados cambiais. Representando com f_t o logaritmo do preço a prazo de uma moeda estrangeira, o chamado preço *forward*, e com s_t o logaritmo do correspondente preço à vista, ou preço *spot*, segundo esta hipótese deve observar-se

$$f_t = \mathrm{E}_t(s_{t+1}),$$

onde E_t representa o valor esperado condicional na informação disponível no final do período t, pois se a relação não se verificar os especuladores podem esperar fazer lucros com as transacções do mercado. Ora, sendo as expectativas dos agentes racionais, deve ter-se

$$s_{t+1} - \mathrm{E}_t(s_{t+1}) = \epsilon_{t+1},$$

com o erro de expectativas $\epsilon_{t+1} \sim I(0)$, com média nula, $\mathrm{E}_t(\epsilon_{t+1}) = 0, \forall t$. Combinando as duas equações tem-se

$$s_{t+1} = f_t + \epsilon_{t+1},$$

que é uma equação de cointegração típica desde que tanto s_{t+1} como f_t sejam $I(1)$. Adicionalmente, neste caso a teoria ou hipótese económica vai bastante longe pois prescreve mesmo o vector de cointegração específico: $[1 \;\; -1]'$ (bastando "puxar" f_t

[2] Esta condição é suficiente para se poder considerar a possibilidade de cointegração? Porquê?

para o lado esquerdo da equação).

Exemplo 3: hipótese da paridade dos poderes de compra (PPP, *purchasing power parity*). Talvez seja boa ideia começar por recordar o índice Big Mac da *Economist* e partir daí para a generalização. Representando com p_t o logaritmo do nível interno de preços e com p_t^* o do nível externo, e com e_t o logaritmo do preço em moeda interna da moeda externa, isto é, a taxa de câmbio, segundo uma versão da PPP deve observar-se

$$e_t = p_t - p_t + u_t,$$

onde já se introduziu um erro para representar os desvios de equilíbrio, que se devem a custos de transporte, fricções nos mercados, direitos alfandegários, etc. . Todavia, como esse erro deve ser estacionário e como e_t, p_t e p_t^* são tipicamente $I(1)$, a equação

$$e_t - p_t + p_t^* = u_t \sim I(0)$$

representa uma relação de cointegração e, também aqui, a teoria económica vai bastante longe pois especifica mesmo o vector de cointegração: $[1 \quad -1 \quad 1]$.

Exemplo 4: função procura de moeda estável. Este é um exemplo mais complexo, não só pelo número de variáveis envolvidas como pelo facto de o nível de integração de algumas delas não ser o simples e comum $I(1)$ mas sim o muito mais raro $I(2)$. Ainda assim, pelo facto de ser um dos casos que tem merecido mais atenção na literatura, com centenas — talvez até milhares — de artigos, bem como pela popularidade da teoria económica em causa, apresenta-se desde já este exemplo.

Represente-se com m_t o logaritmo do *stock* nominal de moeda (por exemplo, M1) e com p_t o logaritmo do nível de preços. É muito frequente que ambas as séries sejam I(2), a primeira pelo facto de se tratar de uma série nominal e a segunda com uma tendência de crescimento (de longo prazo) muito suave; mas se a diferença de ambas (ou seja, o logaritmo do quociente das séries originais) for I(1), tem-se também, como veremos, uma situação de cointegração: uma combinação linear de séries com ordem de integração de 2 dá origem a uma série com uma ordem de integração inferior, $I(1)$, que não é nem mais nem menos que o logaritmo da moeda real.

Ora, a teoria económica estabelece que a procura de moeda real é função do volume de transacções da economia, aproximado pelo PIB logaritmizado, y_t — procura por motivo transacções —, e também de uma taxa de juro, r_t — procura por motivo de especulação.

Tem-se, então

$$\underbrace{m_t - p_t}_{I(1)} = \beta_1 + \beta_2 \, y_t + \beta_3 \, r_t + u_t,$$

que é uma equação de cointegração se y_t e r_t forem ambas $I(1)$ mas $u_t \sim I(0)$. Dada a presença de séries $I(2)$ que cointegram numa $I(1)$ que, por sua vez, cointegra com outras também $I(1)$, fala-se em *multicointegração.*

Muitos outros exemplos poderiam ser apresentados porque a definição de cointegração não é especialmente exigente e, portanto, adequa-se a caracterizar, em linhas gerais, muitos fenómenos económicos. Mesmo algumas velhas teorias económicas foram ressuscitadas com o aparecimento e difusão do conceito de cointegração.

Antes de se generalizar o conceito de cointegração, convém fazer duas observações:

a) como qualquer combinação linear de séries $I(0)$ ainda é (trivialmente) $I(0)$, não existe cointegração entre este tipo de séries;

b) uma equação de regressão espúria entre séries $I(1)$ é uma equação em que todas as séries, incluindo a dos erros, são $I(1)$ e, portanto traduz a ausência de uma relação de cointegração:

$$\underbrace{y_t}_{I(1)} = \alpha + \beta \underbrace{x_t}_{I(1)} + \underbrace{u_t}_{I(1)}.$$

Neste caso não existe, portanto, uma relação de equilíbrio de longo prazo envolvendo y_t e x_t (o que não significa que não possa existir mas requerendo, por exemplo, ainda uma terceira variável $I(1)$, z_t).

3 Generalização

A definição geral de cointegração deve-se a Engle e a Granger (1987) e apresenta-se em seguida.

Definição: considere-se um vector $\mathbf{y_t}$ de m séries ou variáveis: $\mathbf{y_t}' = [y_{1t}\ y_{2t}\ \ldots y_{mt}]'$. Diz-se que ele é cointegrado de ordens (d, b), com d e b inteiros positivos, escrevendo-se $\mathbf{y_t} \sim CI(d, b)$, se:

i. cada componente de $\mathbf{y_t}$ (cada série) for $I(d)$: $y_{it} \sim I(d), i = 1, 2, \ldots, m$;

ii. existir um vector $\boldsymbol{\beta} \neq 0$ tal que

$$\boldsymbol{\beta}'\mathbf{y_t} = \beta_1 y_{1t} + \beta_2 y_{2t} + \ldots + \beta_m y_{mt} = u_t \sim I(d-b),$$

com $d \geq b$; $\boldsymbol{\beta}$ é chamado de vector de cointegração.

Ou seja, a definição exige simplesmente que a combinação linear de séries integradas da mesma ordem dê origem a uma série com uma ordem de integração inferior à delas. Embora sem qualquer relevância que não a de exemplificar, pode ser, por exemplo, $CI(17, 8)$: uma combinação linear de séries $I(17)$ dá origem a uma série $I(9)$.

A exposição auxiliar para a compreensão do conceito e das suas implicações faz-se em seguida através de um conjunto de observações.

a) O caso mais importante em economia é, de longe, o caso $CI(1, 1)$: pela frequência com que se encontram séries $I(1)$ mas, também, pela tradução das relações económicas em termos de relações de equilíbrio de longo prazo.

b) Quando estão envolvidas apenas duas séries $I(1)$ existe um único vector de cointegração linearmente independente.

A demonstração desta proposição faz-se por absurdo. Comece por se representar o vector de cointegração com $[1 \;\; \beta_1]'$ onde, sem perda de generalidade, se normalizou relativamente à primeira variável:

$$y_{1t} + \beta_1 y_{2t} = u_{1t} \sim I(0),$$

onde y_{1t} e y_{2t} são ambas $I(1)$. Contraditoriamente com a tese a provar, admita-se a existência de outro vector de cointegração diferente, $[1 \;\; \beta_2]'$, com $\beta_2 \neq \beta_1$ para serem linearmente independentes:

$$y_{1t} + \beta_2 y_{2t} = u_{2t} \sim I(0).$$

Subtraindo membro a membro as duas combinações lineares, tem-se

$$\beta_1 y_{2t} - \beta_2 y_{2t} = \underbrace{u_{1t} - u_{2t}}_{I(0)},$$

pois u_{1t} e u_{2t} são ambas $I(0)$. Mas o membro esquerdo, igual a $(\beta_1 - \beta_2)y_{2t}$ só pode ser $I(0)$ no caso trivial em que $\beta_1 = \beta_2$, contrariamente à hipótese inicial. O segundo vector terá, então, que ser igual ao primeiro.

Bem entendido, qualquer produto de um vector de cointegração por uma constante não nula é ainda vector de cointegração: se $[1 \;\; \beta_1]'$ é vector de cointegração, então também $[k \;\; k\beta_1]'$ o é pois

$$k\, y_{1t} + k\, \beta_1 y_{2t} = k\, u_t \sim I(0).$$

O número de vectores de cointegração é, portanto, infinito. O que é apenas um é o número de vectores de cointegração linearmente independentes.

c) Todavia, dadas m séries $I(1)$ podem existir, no máximo, $m-1$ vectores de cointegração linearmente independentes. Prove-se que não podem existir m.

Represente-se com $\boldsymbol{\beta}_i$ cada vector de cointegração, $\boldsymbol{\beta}_i'\mathbf{y}_t \sim I(0)$, e construa-se a matriz cujas colunas são os vectores de cointegração, em número de r:

$$\mathbf{B} = [\boldsymbol{\beta}_1\; \boldsymbol{\beta}_2 \;\; \ldots \;\; \boldsymbol{\beta}_r]\,.$$

Como a demonstração também é por absurdo, comece por se admitir, contraditoriamente com o resultado a provar, que existem m vectores de cointegração linearmente independentes. Então, a matriz $\mathbf{B}$ é $(m \times m)$ e regular e $\mathbf{B}'\mathbf{y}_t$ é um vector de m variáveis $I(0)$. Ora, nessas circunstâncias, também

$$(\mathbf{B}')^{-1} \underbrace{\mathbf{B}'\mathbf{y}_t}_{I(0)} = \mathbf{I}\,\mathbf{y}_t = \mathbf{y}_t$$

teria que ser, afinal, um vector de séries $I(0)$, contrariando a hipótese inicial. Não podem, portanto, existir m vectores de cointegração linearmente independentes.

d) Ao número de vectores de cointegração linearmente independentes chama-se *característica de cointegração* e representa-se com r: $r \leq m-1$.

A possibilidade da existência de vários vectores de cointegração coloca um problema de interpretação económica, isto é, de identificação: uma equação de cointegração pode ser considerada uma equação de forma reduzida; ora, como se podem identificar relações de comportamento, estruturais, com significado económico, a partir de simples relações de forma reduzida?

Por exemplo, se aumentarmos o conjunto de variáveis do exemplo da função de procura de moeda com a da inflação, poder-se-á passar a ter duas relações de cointegração. Neste caso, desde que a segunda envolva apenas a inflação e a taxa de juro, a existência de dois vectores de cointegração tem significado económico e é interpretável. Mas nem sempre será assim, o que poderá colocar problemas adicionais à análise. Esta é uma questão muito relevante mas que, a não ser nos casos mais simples, ultrapassa o âmbito deste texto.

Na verdade, a possibilidade de $r > 1$ surge apenas no contexto da modelação multi-equacional ou de sistema, que não será abordado aqui. O estudo a prosseguir far-se-á no contexto uni-equacional e assumindo que existe, no máximo, um único vector de cointegração (linearmente independente).

e) A definição apresentada é a de *cointegração estocástica* pois admite-se que as combinações lineares das variáveis que eliminam as raízes unitárias possam ter tendências lineares. Isto é, cada combinação linear aniquila apenas a tendência estocástica.

O conceito de *cointegração determinística* é mais forte pois exige que os mesmos vectores $\boldsymbol{\beta}_i$ que eliminam as raízes unitárias também eliminam as (eventuais) tendências determinísticas das séries. Formalmente, como se sabe através da decomposição de Beveridge e Nelson e em traços largos, tem-se, para cada série

$$y_{it} = \mathrm{TD}_{it} + z_{it},$$

onde TD representa a tendência determinística, que pode consistir apenas numa constante se a série não tiver deriva ou pode até simplesmente não existir, e z_{it} representa a "função ruído" da série, isto é, a componente estocástica, contendo, designadamente, a tendência estocástica. Ora, deve-se permitir que todas as variáveis tenham tendências determinísticas não nulas, considerando um vector que as contenha

$$\mathbf{TD}_t = \boldsymbol{\gamma}_0 + \boldsymbol{\gamma}_1 t,$$

com $\mathbf{TD}_t$ e $\boldsymbol{\gamma}_0$ e $\boldsymbol{\gamma}_1$ vectores $(m \times 1)$.

Então, a *cointegração determinística* exige não só que $\boldsymbol{\beta}'\mathbf{z}_t \sim I(0)$ mas também que o vector $\boldsymbol{\beta}$ satisfaça a condição $\boldsymbol{\beta}'\mathbf{TD}_t = \text{constante}$. Isto é, que também as tendências determinísticas sejam aniquiladas pelo mesmo vector que aniquila a tendência estocástica; ora esta condição é muito exigente e é raramente satisfeita.

f) Admitindo que a cointegração é "apenas" estocástica, para que a combinação linear tenha média nula — facilitando a interpretação dos erros como sendo de equilíbrio — é frequentemente necessário incluir um termo independente na equação bem como um termo de tendência determinística (e eventualmente também *dummies* sazonais).

Assumindo um único vector de cointegração tem-se, então,

$$u_t = \boldsymbol{\beta}' \mathbf{y}_t - (\gamma_0 + \gamma_1 t),$$

isto é, em termos matriciais:

$$\mathbf{u} = \mathbf{Y}\,\boldsymbol{\beta} - \ \mathbf{D}\,\boldsymbol{\gamma},$$

com $\mathbf{D}$ a representar a matriz das observações dos regressores determinísticos e $\boldsymbol{\gamma}$ o respectivo vector de coeficientes.

g) Saliente-se que pode não haver qualquer relação de cointegração envolvendo apenas duas séries $I(1)$ mas que a inclusão de uma terceira (ou de uma terceira e uma quarta, etc.) pode conduzir a uma relação de cointegração. Por exemplo, no caso da hipótese da paridade dos poderes de compra estão envolvidas três séries.

No resto deste texto serão tratados os dois problemas que requerem resolução:

i) como estimar o vector de cointegração (assumido único)?

ii) Como testar a existência de cointegração?

Note-se que a implementação dos métodos de estimação assume que existe cointegração e, portanto, eles são logicamente posteriores à utilização dos testes. Todavia, por razões de conveniência serão os primeiros a ser abordados.

Os métodos de estimação podem ser classificados como uni-equacionais, por um lado, e de sistema ou multi-equacionais, por outro. Os primeiros assumem que a característica de cointegração é 1 e só permitem estimar um vector de cointegração. Os segundos não só permitem estimar os eventuais vários vectores de cointegração como permitem determinar o seu número, isto é, a característica do espaço de cointegração. Entre estes, destaca-se, pelas suas propriedades, o método de Johansen, que não será abordado aqui. Como é claro do título deste texto, a análise prosseguirá assumindo um único vector de cointegração, a estimar com uma só equação.

4 Estimação do Vector de Cointegração

Abordar-se-ão três métodos de estimação uni-equacionais: SOLS,[3] DOLS e do estimador baseado num modelo dinâmico (ADL).

4.1 Estimação SOLS

Retome-se a relação $u_t = \boldsymbol{\beta}'\mathbf{y}_t - (\gamma_0 + \gamma_1 t)$ e normalize-se o coeficiente da primeira variável, y_{1t}, isto é, faça-se $\beta_1 = 1$:

$$y_{1t} = \mathbf{d_t'}\boldsymbol{\gamma} + \mathbf{y}_{2t}'\boldsymbol{\beta}^* + u_t,$$

com $\boldsymbol{\beta}^* = [-\beta_2^* \ \ldots \ -\beta_m^*]'$, com β_j^*, $j = 2, \ldots, m$ os restantes coeficientes resultantes da normalização. Note-se que esta normalização é, em termos estatísticos, puramente arbitrária, isto é, poder-se-ia ter normalizado de maneira diferente; ou seja, há m maneiras diferentes de normalizar, de escrever a equação de regressão, todas aceitáveis. Em geral, é a teoria económica que indica o regressando e os regressores[4].

O estimador SOLS, sugerido e analisado inicialmente por Engle e Granger, não é mais que o estimador OLS desta equação, estática — *static* OLS. Há boas e más notícias sobre este estimador neste contexto. Comecemos pelas primeiras.

Se as variáveis são cointegradas, então, muito provavelmente, terão sido geradas de forma conjunta e simultânea, donde resulta um problema vulgar de endogeneidade dos regressores. Todavia, ao contrário do que seria de esperar, não decorre daqui qualquer problema de inconsistência do estimador OLS. Para ilustrar esta afirmação considerar-se-á um pequeno exemplo, onde as propriedades apresentadas inicialmente serão muito úteis.

Suponha-se que o PGD é dado pelo sistema

$$\begin{cases} \lambda_1 y_{1t} - y_{2t} = u_{1t}, & \text{com } (1 - \rho_1 L)u_{1t} = \epsilon_{1t}, \ \epsilon_{1t} \sim iid(0, \sigma_1^2), \\ y_{1t} - \lambda_2 y_{2t} = u_{2t}, & \text{com } (1 - \rho_2 L)u_{2t} = \epsilon_{2t}, \ \epsilon_{2t} \sim iid(0, \sigma_2^2), \end{cases}$$

cujo funcionamento é comandado pelos parâmetros ρ_1 e ρ_2:

a) se $|\rho_1| < 1$ e $|\rho_2| < 1$, y_{1t} e y_{2t} são ambas $I(0)$;

[3] A subsecção seguinte baseia-se em Davidson e MacKinnon (1993), capítulo 20.

[4] Apesar desta arbitrariedade, pode haver formas de normalizar melhores que outras; veja-se a subsecção 5.1 mais adiante.

b) se $\rho_1 = \rho_2 = 1$, y_{1t} e y_{2t} são ambas $I(1)$ e não são cointegradas;

c) se, por exemplo, $\rho_1 = 1$ e $|\rho_2| < 1$, $(y_{1t}, y_{2t}) \sim CI(1,1)$ com vector de cointegração $[1 \quad -\lambda_2]'$.

Verifique-se, por exemplo, este último caso, começando por obter a equação da forma reduzida para y_{1t}. Tomando a primeira equação, resolvendo em ordem a y_{2t} e substituindo na segunda equação obtém-se

$$y_{1t} = \underbrace{\frac{1}{1-\lambda_2\lambda_1}u_{2t}}_{I(0)} - \lambda_2 \underbrace{\frac{1}{1-\lambda_2\lambda_1}u_{1t}}_{I(1)} \sim I(1),$$

onde se assumiu que $\lambda_2\lambda_1 \neq 1$. De facto, note-se que se $\rho_1 = 1$ o processo u_{1t} é $I(1)$ pois é um passeio aleatório; e como $|\rho_2| < 1$, $u_{2t} \sim I(0)$ pois segue um AR(1) estacionário. Então, por ser uma combinação linear de um processo $I(1)$ com um $I(0)$, o processo y_{1t} é $I(1)$.

Voltando a substituir na equação de y_{2t} tem-se

$$y_{2t} = \underbrace{\frac{\lambda_1}{1-\lambda_2\lambda_1}u_{2t}}_{I(0)} - \underbrace{\frac{1}{1-\lambda_2\lambda_1}u_{1t}}_{I(1)} \sim I(1).$$

Voltando a considerar a segunda equação do sistema, observa-se agora que ela é uma equação de cointegração com o vector de cointegração referido. Escrevendo-a sob a forma de equação de regressão:

$$\begin{aligned} y_{1t} &= \lambda_2 y_{2t} + u_{2t} \\ &= \lambda_2 y_{2t} + \rho_2 u_{2,t-1} + \epsilon_{2t} \\ &= \lambda_2 y_{2t} + \rho_2 (y_{1,t-1} - \lambda_2 y_{2,t-1}) + \epsilon_{2t}, \end{aligned}$$

onde a última equação volta a resultar da segunda equação do sistema. Esta equação permite verificar que o erro da equação estática de cointegração (u_{2t}) contém dois componentes, ambos fortemente correlacionados com o regressor: o primeiro por ser um desfasamento de uma variável que está cointegrada com ele e o segundo por ser um desfasamento dele próprio, que é uma variável $I(1)$ e, portanto, positiva e fortemente autocorrelacionada. Ou seja, o regressor da equação estática é endógeno. Todavia, como veremos em breve, a existência de cointegração, incluindo a natureza do regressor, $I(1)$, é suficiente para garantir que, neste caso, o problema da endogeneidade desaparece assintoticamente, não afectando a consistência do estimador OLS. Assintoticamente, a variabilidade do regressor (y_{2t}) torna negligenciável o problema da endogeneidade; de resto, pode desde já adiantar-se que, no caso de cointegração, o mesmo acontece mesmo em situações clássicas de simultaneidade e de erros de medição.

Uma suspeita de inconsistência pode também inspirar-se nos casos das regressões espúrias envolvendo variáveis $I(1)$. No entanto, a existência de cointegração afasta esse risco e, pelo contrário, o OLS tenderá a seleccionar o verdadeiro valor do parâmetro ou um valor próximo desse com frequência ainda maior que a usual. Com efeito, intuitivamente, recorde-se que o estimador OLS minimiza a variância estimada dos erros e, para qualquer outro valor de λ_2, os erros serão $I(1)$ e, portanto, têm uma variância que tende para ∞ quando $T \to \infty$.

Na realidade, com cointegração o estimador OLS é até super-consistente: em vez de convergir para o verdadeiro valor dos parâmetros à taxa $\sqrt{T}$, como no caso usual de contextos estacionários, ou seja, em vez de ser $O_p(T^{-1/2})$, a convergência verifica-se à taxa T, o estimador é $O_p(T^{-1})$.

Para verificar e ilustrar esta afirmação no caso do exemplo, escreva-se o estimador OLS em termos do parâmetro a estimar e do erro de amostragem:

$$\widehat{\lambda}_2 = \lambda_2 + \frac{\sum y_{2t}u_{2t}}{\sum y_{2t}^2},$$

onde o erro de amostragem é a concretização da usual expressão $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}$ para o caso presente. Ora o numerador é a soma de produtos de uma série I(1) por uma I(0), que diverge à velocidade ou taxa T, é $O_p(T)$; e o denominador é uma soma de quadrados de uma série I(1), que também diverge mas à velocidade T^2, ou seja, é $O_p(T^2)$. Por conseguinte, como é o quociente de uma magnitude $O_p(T)$ por uma $O_p(T^2)$, $(\widehat{\lambda}_2 - \lambda_2)$ é $O_p(T^{-1})$, ou seja, para obter uma distribuição limite não degenerada é necessário escalar o erro de amostragem com T, não bastando faze-lo com $\sqrt{T}$, como é usual; é a variável $T(\widehat{\lambda}_2 - \lambda_2)$ que é limitada e que tem uma distribuição definida e não $\sqrt{T}(\widehat{\lambda}_2 - \lambda_2)$: o estimador OLS é consistente à taxa T, é *super-consistente*. Em termos intuitivos, a grande variabilidade do regressor sobrepõe-se à sua correlação com o erro, tornando-a assintoticamente desprezível.

Com uma única excepção, as más notícias são em maior número mas não são particularmente graves. Em primeiro lugar, embora o problema da endogeneidade seja assintoticamente desprezível, em amostras finitas ele pode ser significativo. Ou seja, o estimador OLS de regressões estáticas de cointegração é, em geral, enviesado ($\mathrm{E}(\widehat{\lambda}_2) \neq \lambda_2$) e esse enviesamento pode ser grande.

Por outro lado, também regra geral, o estimador OLS não é assintoticamente eficiente. No caso do exemplo, repare-se que o OLS está a estimar uma equação mal especificada em termos dinâmicos, despreza informação e, portanto, não pode ser eficiente.

Por outro lado, ainda, *em geral* os métodos de inferência assintótica continuam a não ser válidos, tal como no caso das regressões espúrias. Porque acontece isto? Devido a dois problemas bem distintos: a endogeneidade dos regressores e a autocorrelação dos erros, típicas das regressões estáticas, fazem aparecer parâmetros perturbadores nas distribuições assintóticas, que invalidam a utilização das estatísticas de teste habituais. Ou seja, em geral, nem as estatísticas-t nem as Fs podem ser validamente empregues para fazer inferências sobre os valores dos parâmetros de cointegração.

Existe, no entanto, uma excepção interessante a esta regra geral em que:

a) os erros da equação de regressão são normais e *iid*;

b) os regressores da equação de cointegração são estritamente exógenos.

Voltando a considerar o caso geral de um vector de m variáveis, esta é a situação em que se tem um sistema triangular:

$$\begin{cases} y_{1t} = \alpha + \boldsymbol{\beta}'\mathbf{y}_{2t} + u_{1t} \\ \mathbf{y}_{2t} = \mathbf{y}_{2,t-1} + \mathbf{u}_{2t}, \end{cases} \tag{1}$$

com uma determinação recursiva, com o subvector $\mathbf{y}_{2t}$ a ser primeiramente determinado pelas equações inferiores, sob a forma de passeios aleatórios aparentes, e em seguida a determinar y_{1t} através da primeira equação, a equação de cointegração.

Adicionalmente, deve ter-se ainda

$$\begin{bmatrix} u_{1t} \\ \mathbf{u}_{2t} \end{bmatrix} \sim iid\mathcal{N}\left(\begin{bmatrix} 0 \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega}_{22} \end{bmatrix} \right),$$

e em particular

$$\mathrm{E}(u_{1t}\mathbf{u}_{2s}) = \mathbf{0} \Leftrightarrow \mathrm{E}(u_{1t}\Delta\mathbf{y}_{2s}) = \mathbf{0}, \forall t, s.$$

Ou seja, além de erros da equação de cointegração *iid* e normais, os regressores são estritamente exógenos. Neste caso, nem sequer é necessário recorrer à teoria distributiva assintótica pois a teoria exacta é válida: as estatísticas usuais t e F têm distribuições exactas t e F, respectivamente. Mas este resultado não nos deve surpreender: afinal, as hipóteses do modelo clássico são satisfeitas, pelo que os seus resultados são válidos, não existe qualquer novidade. É verdade que as variáveis envolvidas não são estacionárias nem fracamente dependentes, mas o modelo de regressão clássico também não requer nenhuma condição desse tipo.

Por outro lado, se a hipótese de normalidade for relaxada e os erros forem apenas *iid*, não necessariamente normais, as distribuições das estatísticas de teste voltam a ser válidas apenas assintoticamente. Além disso, o OLS é assintoticamente eficiente. Note-se, no entanto, que mesmo esta situação é muito restritiva, pouco passível de ser observada: não só não pode existir qualquer efeito de *feedback* como os erros de uma equação de regressão estática, sem qualquer dinâmica, não podem estar autocorrelacionados.

Em termos muito sumariados, meramente informativos, as soluções que podem ser adoptadas para ultrapassar a ineficiência geral do estimador OLS e para poder efectuar inferências sobre os parâmetros de cointegração são, para além dos que veremos em seguida:

a) o estimador FIML de Johansen, no contexto multi-equacional;

b) o método dos 3 passos de Engle e Yoo, simples mas muito raramente empregue;

c) o estimador FM-OLS (*fully modified*) de Phillips e Hansen, em que são aplicadas duas correcções não paramétricas ao estimador OLS (uma para acomodar o problema da autocorrelação dos erros e a outra para tratar do problema da endogeneidade dos regressores [5]).

E se, dadas as m variáveis $I(1)$ existirem $m-1>1$ vectores de cointegração linearmente independentes (violando uma hipótese crucial desta abordagem)? Nesse caso, o estimador OLS pode estar a estimar uma combinação linear desses $m-1$ vectores da base do espaço de cointegração. Em particular, pode provar-se que, entre os possíveis vectores ou relações de cointegração, o estimador OLS selecciona a relação cujos erros não estão correlacionados com nenhuma combinação linear $I(1)$ das variáveis $y_{2t}, y_{3t}, \ldots, y_{mt}$ (veja-se Hamilton, 1994, p. 590-1).

4.2 Estimador DOLS

O estimador DOLS, *dynamic* OLS, por oposição ao OLS anterior da regressão estática, foi proposto no início da década de 90, de forma separada e independente,

[5] Veja-se também, como sucedâneo do original, Boswijk (1994) ou Hamilton (1993), pp. 613-18.

por Saikkonen, Phillips e Loretan e Stock e Watson. Apesar da sua razoável simplicidade é um estimador com muito boas propriedades e surpreende o facto de não ser empregue mais frequentemente.

Retome-se o sistema de equações anterior, (1), mas suponha-se agora que os regressores $\mathbf{y}_{2t}$ são endógenos, $\mathrm{Cov}(u_{1t}, \mathbf{u}_{2t}) \neq 0$, violando uma hipótese crucial do modelo clássico bem como do modelo para séries estacionárias (e regressores pré-determinados). Considere-se, então, a projecção linear de u_{1t} sobre $\mathbf{u}_{2,t-p}, \mathbf{u}_{2,t-p+1}, \ldots,$ $\mathbf{u}_{2,t-1}, \mathbf{u}_{2t}, \mathbf{u}_{2,t+1}, \ldots \mathbf{u}_{2,t+p}$ e seja $\tilde{u}_t$ o respectivo erro:

$$u_{1t} = \sum_{s=-p}^{p} \boldsymbol{\gamma}_s' \mathbf{u}_{2,t-s} + \tilde{u}_t.$$

Logo, por definição de projecção linear, i.e., por construção, $\tilde{u}_t$ não é correlacionado com $u_{2,t-s}$ para $s = -p, -p+1, \ldots, p$. Mas como $u_{2t} \equiv \Delta y_{2t}$, a equação de cointegração pode agora escrever-se como

$$y_{1t} = \alpha + \boldsymbol{\beta}' \mathbf{y}_{2t} + \sum_{s=-p}^{p} \boldsymbol{\gamma}_s' \boldsymbol{\Delta} \mathbf{y}_{2,t-s} + \tilde{u}_t,$$

onde não existe qualquer correlação entre os regressores e o erro desde que $\mathrm{Cov}(u_{1t},$ $u_{2,t-s}) = 0$ para $|s| > p$, ou seja, desde que não exista correlação entre os erros das equações do sistema afastados mais do que p períodos de tempo, os regressores da equação de cointegração aumentada com os desfasamentos e os valores futuros são, agora, estritamente exógenos; presume-se que a correlação que existia inicialmente é totalmente capturada pelos *lags* e *leads* introduzidos.

Desta forma, a correcção paramétrica ao OLS elimina assintoticamente o efeito da endogeneidade sobre a distribuição do estimador OLS: o que é necessário é acrescentar à regressão estática avanços e recuos (desfasamentos) das primeiras diferenças dos regressores. Em termos práticos, a truncatura ou escolha de p pode basear-se na teoria da distribuição assintótica para eliminar regressores não significativos, empregando uma estratégia de modelação do geral para o particular (GTS): iniciar o processo com um número de avanços e desfasamentos que pareça suficiente para acomodar qualquer correlação entre os regressores originais e os erros e "testar para baixo", eliminando regressores em primeiras diferenças que se mostrem irrelevantes.

Se $\tilde{u}_t$ não for autocorrelacionado, o estimador DOLS é assintoticamente eficiente e as estatísticas-t e F sobre os parâmetros de cointegração são válidas assintoticamente. Frequentemente, a inclusão dos *leads* e *lags* é suficiente para "branquear" os erros: a "dinamização" do modelo através dessa introdução remove a autocorrelação dos erros originais e não é necessário proceder a qualquer correcção adicional.

Todavia, se os $\tilde{u}_t$s forem autocorrelacionados, são necessárias correcções adicionais. Aqui os vários métodos divergem, podendo a correcção adicional ser paramétrica ou não. Por exemplo, Phillips e Loretan sugerem a inclusão de desfasamentos do erro de equilíbrio na equação de cointegração, de forma semelhante à da primeira equação do exemplo de sistema que vimos na subsecção anterior:

$$y_{1t} = \alpha + \boldsymbol{\beta}'\mathbf{y}_{2t} + \sum_{s=-p}^{p} \boldsymbol{\gamma}_s' \boldsymbol{\Delta}\mathbf{y}_{2,t-s} + \sum_{s=1}^{p} \theta_s (y_{1,t-s} - \boldsymbol{\beta}'\mathbf{y}_{2,t-s}) + \epsilon_t,$$

que requer a estimação por NLLS. Uma via diferente e mais simples consiste na correcção das estatísticas de teste que emprega uma estimativa paramétrica da variância de longo prazo dos erros da equação dinâmica (veja-se, por exemplo, Hamilton, 1994, pp. 608-10 ou Hayashi, 2000, pp. 655-7).

4.3 Um pequeno exemplo

Para ilustrar os métodos de estimação anteriores retome-se o exemplo da função consumo para a economia portuguesa, com dados de 1965 a 1995, onde, recorde-se, LCP representa o logaritmo do consumo privado, LRD o do rendimento disponível, INF a taxa de inflação e LSR o de um índice de salários reais. Assumindo a existência de cointegração, a estimação SOLS produz, como estimativa da relação de equilíbrio de longo prazo

$$\widehat{LCP} = 0.099 + 0.023A75 + 0.935LRD + 0.192LSR - 0.004INF,$$

onde $A75$ representa uma variável artificial de impulso, com valor unitário apenas no ano de 1975, e onde se omitiram os erros-padrão usuais pois as estatísticas-t não deverão ser válidas não só devido à provável presença de autocorrelação nos erros — fortemente indicada pelas estatísticas usuais — mas também devido à muito plausível endogeneidade dos regressores e, em particular, de LRD.

Continuando a assumir a existência de cointegração, o processo de estimação DOLS foi iniciado com um modelo com desfasamentos (passados e futuros) até à ordem 2, não só pelo facto de os dados serem anuais mas também porque a dimensão da amostra é pequena. Com simplificações sucessivas chegou-se ao modelo:

$$\begin{aligned}\widehat{LCP}_t = \; & -0.037 + 0.105A75_t + 0.976LRD_t - 0.012INF_t + 0.160LSR_t + 0.427\Delta LRD_{t-1} \\ & +0.002\Delta INF_t - 0.007\Delta INF_{t+1} - 0.009\Delta INF_{t+2} - 0.714\Delta LSR_t - 0.619\Delta LSR_{t+2},\end{aligned}$$

cujos erros-padrão dos coeficientes da relação de cointegração são, respectivamente, 0.124, 0.030, 0.021, 0.0015 e 0.062. Desta forma, a constante não é estatisticamente significativa aos níveis usuais; todavia, além de habitual, a sua presença é importante para assegurar a robustez das previsões. Refira-se que também o regressor ΔINF_t não é estatisticamente significativo aos níveis usuais (valor-p de 0.106); a sua presença, contudo, parece importante para assegurar a ausência de quaisquer sintomas de autocorrelação dos erros; como estes são completamente inexistentes (valores-p da BG(1) e BG(3) iguais, respectivamente a 0.867 e 0.143), a validade assintótica dos métodos de inferência usuais parece justificar-se. Assim, por exemplo, o teste de $H_0 : \beta_3 = 1$ *vs.* $H_1 : \beta_3 \neq 1$, de que a elasticidade consumo-rendimento é unitária, pode basear-se na estatística-t usual, com valor de $\frac{0.976-1}{0.021} = -1.143$, insignificante aos níveis usuais.

Pergunta-se: os erros da equação estimada pelo DOLS podem ser empregues para analisar a existência de cointegração? Não, porque os erros da equação já são distintos dos erros da equação de cointegração, estática; não são a entidade aqui representada com u_t. De resto, o procedimento DOLS assume a existência de cointegração, não serve para testar essa existência.

4.4 Estimador do modelo ADL — revisão

Recorde-se que os modelos ADL também permitem obter facilmente estimativas dos multiplicadores de longo prazo, que não são nem mais nem menos que os parâmetros de cointegração pois estes são os coeficientes das relações de equilíbrio de longo prazo. Recorde-se, também, que a forma preferida para esses modelos é a de Bardsen.

Assim, considerando um ADL(r,s) com k variáveis explicativas assumidas como (fracamente) exógenas, tem-se

$$\Delta y_t = \mu - A(1)y_{t-1} + \sum_{i=1}^{r-1} \delta_i \Delta y_{t-1} + \mathbf{B}(1)'\mathbf{x}_{t-1} + \sum_{j=0}^{s-1} \boldsymbol{\gamma}_i' \boldsymbol{\Delta}\mathbf{x}_{t-j} + \epsilon_t,$$

onde $A(L)$ representa o polinómio autoregressivo e $\mathbf{B}(L)$ o vector de polinómios de desfasamentos distribuídos das variáveis exógenas ($B_j(L) = \beta_{0j} + \beta_{1j}L + \beta_{2j}L^2 + \ldots + \beta_{s_j,j}L^{s_j}, j = 1, \ldots, k$).

Recorde-se também que, excluindo a constante, os parâmetros de cointegração

Quadro 1. Estimativas dos parâmetros de cointegração

variável	OLS	DOLS	ADL
C	0.099	-0.037	0.183
$A75$	0.023	0.105	—
LRD	0.935	0.976	0.914
LSR	0.192	0.160	0.358
INF	-0.004	-0.012	-0.0003

são dados por

$$\begin{aligned} \boldsymbol{\lambda} = & \ [\boldsymbol{\lambda}_1 \ \ \boldsymbol{\lambda}_2 \ldots \ \boldsymbol{\lambda}_k]' \\ = & \ A(1)^{-1}\mathbf{B}(1)'. \end{aligned}$$

Se se pretender incluir também a constante na relação de cointegração, basta fazer $\lambda_0 = \mu/A(1)$.

4.5 De regresso ao exemplo

Regressando ao exemplo empírico, recorde-se que se obteve, iniciando a modelação com um ADL(3,3) e sem incluir qualquer *dummy*[6]:

$$\widehat{LCP} = 0.183 + 0.914LRD + 0.358LSR - 0.00034INF.$$

Em resumo, os 3 métodos de estimação produzem as estimativas dos coeficientes de cointegração que se apresentam no Quadro 1.

Note-se a concordância do sinal do coeficiente da inflação produzida pelos diferentes métodos, ainda assim sem se poder considerar claramente que a dúvida do efeito de longo prazo desta variável sobre o consumo fica esclarecida empiricamente pois as magnitudes obtidas são sempre muito baixas. De resto e em geral, para além dos sinais, também as magnitudes dos coeficientes estimados são bastante plausíveis; exceptua-se a estimativa do coeficiente de LRD obtida com o DOLS, que parece muito elevada, atendendo até ao período amostral. No entanto, repare-se

[6] Nesta modelação a *dummy* de impulso para 1975 e a variável LSR tendem-se a excluir-se mutuamente, o que não constitui qualquer surpresa pois uma das principais marcas de 1975 foi precisamente o forte crescimento dos salários. Quando se inclui essa *dummy*, obtém-se:

$$\widehat{LCP} = 0.175 + 0.221A75 + 0.958LRD - 0.0099INF.$$

também que é com este método de estimação que a influência da variável de índice salarial sobre o consumo aparece como mais esbatida. Pelo contrário, quando este efeito aparece como mais importante, na estimação baseada no modelo dinâmico, é também quando a elasticidade de longo prazo consumo-rendimento é menor.

5 Testes de Cointegração

Considere-se de novo a, agora apenas potencial equação de cointegração, contendo eventualmente regressores determinísticos[7]:

$$y_{1t} = \mathbf{d}_t'\boldsymbol{\gamma} + \mathbf{y}_{2t}'\boldsymbol{\beta} + u_t.$$

Se as variáveis são cointegradas, então u_t representa o erro da relação de equilíbrio de longo prazo e, por conseguinte, $u_t \sim I(0)$; pelo contrário, se não são cointegradas, então $u_t \sim I(1)$.

Por conseguinte, um possível procedimento de teste consiste em testar

$$\begin{aligned} &H_0 : u_t \sim I(1) \Leftrightarrow \text{não cointegração}, \\ \text{vs. } &H_1 : u_t \sim I(0) \Leftrightarrow \text{cointegração}. \end{aligned}$$

Note-se que, empregando a teoria económica, os papéis das duas hipóteses deveriam ser invertidos. Todavia, como os testes mais populares de raiz unitária são aqueles em que esta hipótese é tomada como nula e como há abundância destes testes, são eles os mais comuns. Se o vector de cointegração for conhecido, isto é, dado também pela teoria económica — como nos casos das hipóteses de eficiência dos mercados ou da paridade dos poderes de compra, por exemplo —, então ele é imposto aos dados e pode efectuar-se facilmente um teste de cointegração fazendo um vulgar teste de raiz unitária sobre u_t.

5.1 Testes de Engle-Granger

Em geral, todavia, o vector de cointegração não é conhecido, pois raramente a teoria económica o especifica. Por conseguinte, u_t não é observável e os testes que se referiram há pouco não podem ser feitos. A sugestão de Engle e de Granger para ultrapassar este problema consistiu em empregar a solução usual, substituindo os erros u_t não observados pelos resíduos OLS, e_t (ou $\widehat{u}_t$). Ou seja, num primeiro passo

[7] Trata-se da equação inicial da subsecção 4.1 mas deixando cair os asteriscos dos coeficientes.

estima-se a potencial equação de cointegração, para a obtenção desses resíduos e, em seguida, efectua-se o teste de

$$H_0 : e_t \sim I(1) \Leftrightarrow \text{não cointegração,}$$
$$\text{vs. } H_1 : e_t \sim I(0) \Leftrightarrow \text{cointegração,}$$

sucedâneo do teste acima, empregando uma regressão de Dickey-Fuller.

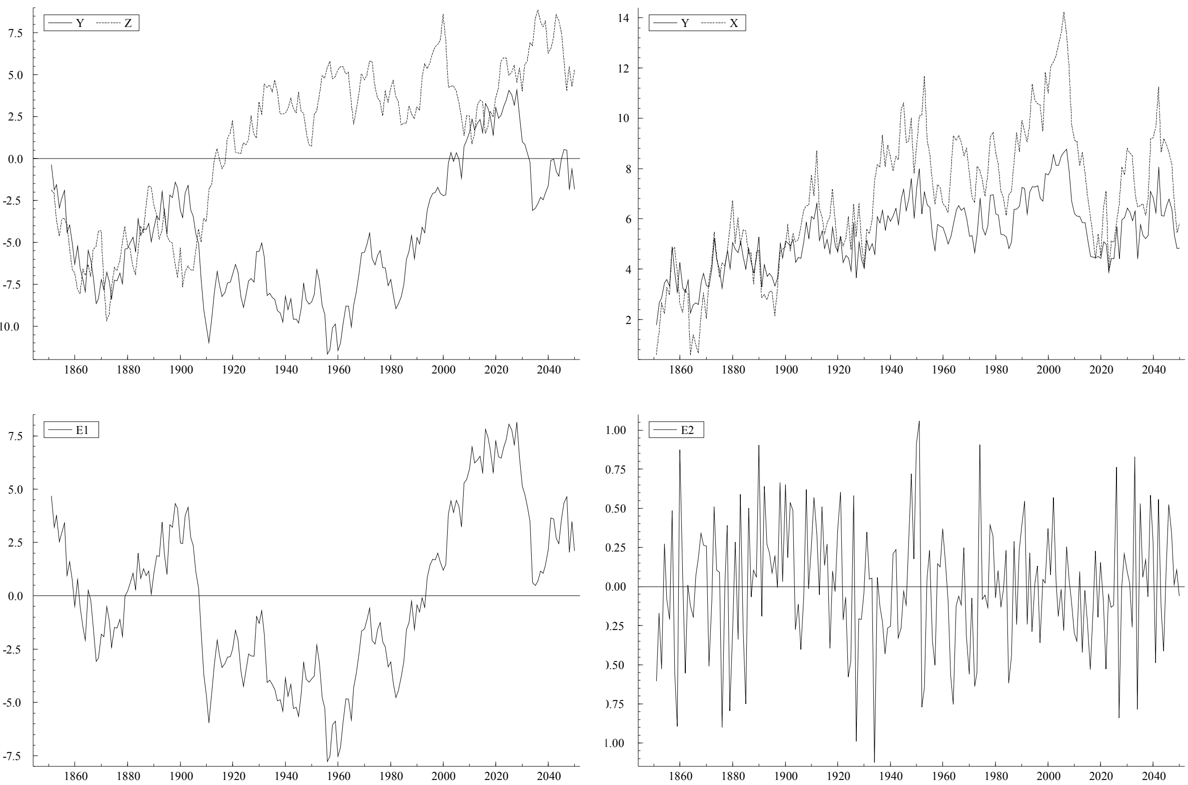

Figura 1: a) no painel superior esquerdo, as duas séries não são cointegradas, são dois passeios aleatórios independentes; no painel inferior, e_{1t} representa a série dos resíduos OLS da regressão de y_t sobre z_t. b) no painel superior direito as duas séries são cointegradas e, por baixo, a série e_{2t} é a série dos resíduos OLS da regressão de y_t sobre x_t.

A ideia básica destes testes para o caso de apenas duas séries é ilustrada pela figura 1. No painel do lado esquerdo, as séries y_t e z_t são duas séries $I(1)$ não cointegradas: trata-se de dois passeios aleatórios independentes, gerados com base em choques independentes; embora por vezes, sobre períodos de tempo curtos, pareçam acompanhar-se, na realidade não covariam de forma sistemática sobre prazos um pouco longos. Por exemplo, numa fase inicial parecem acompanhar-se, primeiro decrescendo e depois crescendo de forma aparentemente relacionada mas, a partir de certa altura, passam a divergir de forma bastante clara, em muitos períodos consecutivos. Traduzindo esta ausência de uma relação de longo prazo, no painel inferior esquerdo a série dos resíduos OLS da regressão de uma série sobre a outra tem um aspecto gráfico típico de uma série $I(1)$: embora a presença de uma constante na regressão imponha aos resíduos a média nula, eles raramente regressam a essa média, cruzando a linha do eixo das abcissas, e o seu comportamento é muito suave; a partir de certa altura também parecem necessitar de um intervalo de variação maior que o inicial e a partir de 1980 esse intervalo é claramente maior que o das primeiras observações.

Já as duas séries do painel superior direito são ($I(1)$ e) cointegradas: embora por vezes não sigam o padrão usual oscilando em sentidos opostos e se cruzem até por diversas vezes devido a oscilações discordantes, em geral elas tendem a variar de forma conjunta, acompanhando-se no longo prazo. A reflectir este comportamento, a série dos resíduos OLS da regressão de uma (y_t, um outro y_t distinto do do lado esquerdo) sobre a outra (x_t) tem um aspecto visual claramente estacionário: muito oscilante, a cruzar frequentemente o eixo das abcissas, sem qualquer tendência de divergência apenas para valores negativos ou positivos sobre períodos relativamente longos, variando sempre no interior de uma banda com largura aproximadamente constante, etc. .

O que já foi apresentado como um aspecto muito positivo do OLS tem agora o seu reverso: como o estimador OLS escolhe o vector de cointegração de forma a minimizar a variância dos resíduos, força-os a parecerem o mais estacionários possível (mesmo quando não o são). Desta forma, a distribuição de DF já não é válida para efectuar os testes de DF sobre os resíduos: os valores críticos da distribuição de DF terão que ser substituídos por outros, ainda mais elevados em valor absoluto. Como os primeiros valores críticos das distribuições, obtidos por simulação, foram apresentados por Engle e Granger, a distribuição tem o nome de Engle-Granger (EG).

Em geral, as regressões de teste são do tipo

$$\Delta e_t = \phi e_{t-1} + \sum_{i=1}^{k} \gamma_i \Delta e_{t-i} + \epsilon_t,$$

com e_t a representar os resíduos OLS da potencial equação de cointegração, $\phi = \rho - 1$ e com k suficientemente grande (mas não em excesso) para "branquear" os erros.

Ou seja, note-se que, em geral, não se inclui a constante nesta regressão pois ela terá já sido incluída na potencial equação de cointegração, garantindo a nulidade da média de e_t. Também não se inclui um termo de tendência porque as variáveis incluídas nessa regressão já o devem incluir, de forma implícita ou explícita.

Note-se ainda que as distribuições dependem aindo do número de variáveis envolvidas: quanto mais variáveis mais os resíduos são "maquilhados" ou "mascarados" pelo OLS e, portanto, mais exigentes terão que ser os valores críticos. Estes são apresentados no Quadro 2 e, em geral, são os de Phillips e Ouliaris (1990), seguindo a forma de apresentação de Hayashi (2000).

Há 3 casos a considerar:

1. $\mathrm{E}(\Delta y_{1t}) = 0$ e $\mathrm{E}(\Delta \mathbf{y}_{2t}) = 0$, isto é, nenhuma das séries tem deriva. Este é o caso mais simples e os valores críticos são os do painel superior do quadro.

Quadro 2. Valores críticos assimptóticos dos testes EG (DF sobre os resíduos)

g	1%	5%	10%
	caso A: regressores sem deriva		
1	−3.96	−3.37	−3.07
2	−4.31	−3.77	−3.45
3	−4.73	−4.11	−3.83
4	−5.07	−4.45	−4.16
5	−5.28	−4.71	−4.43
	caso B: regressores com deriva		
1	−3.96	−3.41	−3.13
2	−4.36	−3.80	−3.52
3	−4.65	−4.16	−3.84
4	−5.04	−4.49	−4.20
5	−5.36	−4.74	−4.46

Fontes: Phillips e Ouliaris (1990), quadros IIb) e IIc), p. 190, excepto a primeira linha do painel inferior, retirada de Fuller (1996), quadro 10.A.2. g representa o número de regressores $I(1)$.

2. $\mathrm{E}(\Delta \mathbf{y}_{2t}) \neq 0$ e $\mathrm{E}(\Delta y_{1t})$ pode ser 0 ou não, isto é, pelo menos um dos regressores tem deriva (e a variável dependente pode ter ou não). Mesmo que sejam vários os regressores com deriva, as suas tendências lineares podem ser combinadas numa só e a regressão de teste pode escrever-se como a regressão de y_{1t} sobre uma constante, $g-1$ regressores sem deriva e um regressor com deriva. Ora, tal

como acontece graficamente, também o comportamento estatístico assintótico dos regressores com deriva é dominado pela sua deriva, o que significa que o comportamento assintótico do regressor $I(1)$ com deriva é assintoticamente equivalente ao da tendência linear (t). Assim, também os resíduos assintoticamente se comportam de forma idêntica aos de uma regressão com constante, $g-1$ regressores sem deriva e o tempo. Por exemplo, se $g=1$ apenas, a regressão é assintoticamente equivalente a uma regressão da variável dependente sobre a constante e a tendência linear, ou seja, a uma simples regressão de DF. Neste caso a tabela a usar é a do painel inferior do quadro (e confronte-se o caso anterior com o dos valores críticos da distribuição de DF).

3. $\mathrm{E}(\Delta\mathbf{y}_{2t})=0$ e $\mathrm{E}(\Delta y_{1t})\neq 0$, ou seja, y_{1t} tem deriva mas $\mathbf{y}_{2t}$ não, como por exemplo, se se tentar explicar o comportamento de uma série de investimento que tende a crescer ao longo do tempo com base numa taxa de juro. Neste caso, para remover dos erros a tendência é necessário incluir o termo em t na regressão. Então, a estatística obtida da regressão com constante, tendência e g regressores $I(1)$ tem a distribuição dada pelo painel inferior do quadro mas com $g+1$ regressores; por exemplo, se $g=3$ o valor crítico a 5% é -4.49.

Este último caso, em particular, leva a colocar a questão: deve incluir-se t na regressão quando tal se justifique? É que mesmo nos casos 1 e 2 pode incluir-se o termo de tendência na regressão desde que se empregue o painel inferir do quadro com $g+1$ regressores. Esta prática teria a vantagem de se empregarem sempre os mesmo valores críticos, independentemente da natureza das séries. Todavia, estudos de simulação indicam que a inclusão de t nas regressões, mesmo quando se justifica, lhes retira potência, pelo que deve ser evitada (nas regressões de teste, entenda-se).

Relativamente à determinação da ordem da autoregressão, os (agora chamados) testes AEG não foram objecto de estudos comparáveis aos testes ADF e, portanto, não existe qualquer indicação especial. Convirá dar prioridade aos métodos GTS t-sig e à utilização do critério AIC.

Note-se ainda que, embora assintoticamente equivalentes, os resultados de teste podem, em amostras finitas, ser sensíveis à variável escolhida como dependente, isto é, à normalização adoptada. A este respeito, Ng e Perron (1997) mostram que se a primeira diferença de uma série tem autocorrelação negativa, as estimativas baseadas na equação que usa essa série como variável dependente têm sempre melhores propriedades que as que são normalizadas com outra variável como dependente. Assim, dado que os testes de raiz unitária vão tender a rejeitar mais frequentemente a nula para essas séries, recomendam que seja a variável "menos integrada" que deve ser usada como variável dependente. Intuitivamente esta recomendação faz sentido pois é bem sabido que convém que os regressores tenham muita variabilidade para identificar os seus efeitos com precisão; ora as variáveis que apresentam maior variabilidade deverão ser as "mais integradas".

Refira-se, finalmente, que se os testes AEG não fornecerem evidência estatística de cointegração, isso equivale a uma indicação de que a regressão estática em causa parece ser, afinal, uma regressão espúria (embora possa não o ser eventualmente no sentido integral do termo).

5.2 Revisitando o exemplo

A questão a responder agora é: mas afinal existe cointegração entre LCP_t, LRD_t, LSR_t e INF_t (incluindo também uma constante e a *dummy* na regressão)?

Empregando regressões de teste AEG, iniciando a análise com $k_{MAX} = 3$ e usando o método GTS para determinação do número óptimo de desfasamentos chega-se a $\widehat{k} = 0$ e obtém-se $t_\phi = AEG(0) = -3.08$ (denotando $\rho - 1$ com ϕ). Como estamos no caso 2 pois LRD_t tem deriva e como $g = 3$, o valor crítico assintótico a 5% é -4.16, isto é,

$$RC = \{t_\phi : t_\phi < -4.16\}.$$

Assim, está-se um pouco longe de obter evidência favorável à existência de cointegração. Quando se trocam os papéis das variáveis e se emprega LRD_t como variável dependente, volta-se a chegar a $\widehat{k} = 0$ e $t_\phi = AEG(0) = -3.29$, que é um pouco melhor mas ainda claramente insuficiente para conseguir rejeitar a nula de não cointegração. Parece, assim que, embora não espúria no sentido negativo usual, a regressão estimada anteriormente não tem valor, não corresponde a uma relação de equilíbrio de longo prazo. Pode ser, no entanto, que a reduzida dimensão da amostra, a par com a natureza do teste, estejam a impedir a obtenção de uma conclusão positiva.

6 Cointegração e MCE

Como vimos, mesmo no caso mais favorável, o da existência de cointegração, *em geral*, os métodos usuais de inferência, assentes nas estatísticas t e F, não são válidos. Uma possível solução, sugerida há muito tempo pelos analistas de séries temporais, consiste em empregar as primeiras diferenças das séries para permitir a aplicação dos métodos convencionais. O problema com esta sugestão é que, no caso de as variáveis serem cointegradas, se despreza informação sobre o longo prazo, isto é, não é usada a informação sobre a existência de uma relação de equilíbrio de longo prazo, que envolve os níveis das variáveis, não as suas primeiras diferenças. Ou seja, trabalha-se apenas com relações de dinâmica de curto prazo que não incorporam, ignoram, a restrição da existência de um equilíbrio de longo prazo.

6.1 O Teorema de Representação de Granger

Ora, uma forma de ultrapassar este problema consiste em recorrer ao modelo de correcção de erros (MCE), que concilia o ajustamento dinâmico de curto prazo com a incorporação da relação de equilíbrio de longo prazo. Esse recurso é justificado pelo *Teorema de Representação de Granger*, que será apresentado sob uma forma muito ligeira e, em termos formais e para simplificar a exposição, apenas para o caso de duas variáveis.

Se $(y_t, x_t) \sim CI(1,1)$ com parâmetro de cointegração λ, então as variáveis podem ser representadas através do MCE vectorial seguinte:

$$\begin{cases} \Delta y_t = \delta_1 + \alpha_1(y_{t-1} - \lambda x_{t-1}) + \sum \gamma_{1i}\Delta y_{t-i} + \sum \delta_{1j}\Delta x_{t-j} + \epsilon_{1t} \\ \Delta x_t = \delta_2 + \alpha_2(y_{t-1} - \lambda x_{t-1}) + \sum \gamma_{2i}\Delta y_{t-i} + \sum \delta_{2j}\Delta x_{t-j} + \epsilon_{2t} \end{cases}$$

onde pelo menos um dos αs é diferente de zero ($\alpha_1 \neq 0 \vee \alpha_2 \neq 0$). O inverso também é verdadeiro: se as variáveis são $I(1)$ e admitem uma representação de MCE vectorial, então são cointegradas.

Para compreender melhor o significado e as implicações deste resultado apresentam-se várias observações.

1. O MCE vectorial é um modelo VAR nas primeiras diferenças das variáveis aumentado com o termo de correcção de erros, isto é, na linha da crítica feita anteriormente, o modelo só nas primeiras diferenças das variáveis está mal especificado pois omite o termo de correcção de erros (ou de desequilíbrios).

2. Como, apesar de o modelo incluir algumas variáveis em níveis, todas as variáveis envolvidas se encontram sob a forma $I(0)$, os métodos de inferência usuais não deverão encontrar dificuldades especiais; aliás, a este respeito, note-se

ainda que a notação empregue para os erros das duas equações indica que estes deverão ser processos ruído branco e, portanto, homoscedásticos e não autocorrelacionados.

3. Omitem-se propositadamente os índices dos somatórios pois não existe qualquer restrição teórica neste aspecto; os desfasamentos serão os suficientes para garantir que os erros são ruído branco[8].

4. Muito frequentemente, o interesse reside apenas na primeira equação, que se pode considerar que representa a distribuição de y_t condicional em x_t e nos desfasamentos de ambas. Ora, os métodos de estimação e inferência baseados apenas nessa equação, desprezando a segunda, são válidos e eficientes desde que seja satisfeita a restrição $\alpha_2 = 0$. Diz-se então, neste caso, que x_t é fracamente exógena[9] para os parâmetros (da primeira equação) do MCE e, nesse caso, a segunda equação é vista como representando a distribuição marginal de x_t.

   Todavia, se $\alpha_2 \neq 0$ a estimação isolada da primeira equação não é eficiente pois o parâmetro λ encontra-se presente em ambas, isto é, existe uma restrição inter-equações que deve ser respeitada para estimar o modelo de forma eficiente. Então, a análise deveria ser iniciada estimando de forma conjunta as duas equações e testando $H_0 : \alpha_2 = 0$, isto é, fazendo um teste de exogeneidade fraca. Ou seja, embora pareça absurdo, deveria iniciar-se a análise estimando de forma conjunta as duas equações para saber se essa estimação pode ser dispensada.

   Na prática, dada a restrição aos métodos uni-equacionais, assumir-se-á que $\alpha_2 = 0$ e concentrar-se-á a atenção no MCE condicional.

[8] Contudo, o número de desfasamentos não seria livre se se tivesse partido de uma certa ordem para o modelo VAR inicial, nos níveis das variáveis.

[9] Seja $f(y, x; \theta)$ a densidade conjunta das variáveis Y e X. Ela pode ser sempre factorizada como

$$f(y, x; \theta) = g(y|x; \phi_1)h(x; \phi_2).$$

Suponha-se que o interesse reside num certo vector de parâmetros, ψ. Então: a) se ψ é função apenas de ϕ_1, $\psi = l(\phi_1)$; b) e se ϕ_1 e ϕ_2 são de "variação livre", então diz-se que X é fracamente exógena para ψ e a inferência baseada em $g(y|x)$ é válida e completamente eficiente. Ainda relativamente à noção de "variação livre", considerem-se os espaços dos parâmetros ϕ_1 e ϕ_2: $\phi_1 \in \Phi_1$ e $\phi_2 \in \Phi_2$. Ser de variação livre significa que $(\phi_1, \phi_2) \in \Phi_1 \times \Phi_2$, isto é, o espaço Φ_1 não é função de ϕ_2 e o espaço Φ_2 não é função de ϕ_1, ou seja, o conhecimento sobre o valor de um parâmetro não fornece informação sobre o domínio de potenciais valores do outro.

6.2 O teste t-MCE

Para além de abrir o caminho para os modelos de correcção de erros, o Teorema de Representação de Granger também sugere uma via alternativa para testar a existência de cointegração: testando a presença de um mecanismo de correcção de erros, com o teste t-MCE, já referido aquando do estudo dos modelos dinâmicos. No contexto da primeira equação do sistema, trata-se de testar

$$H_1 : \alpha_1 = 0 \quad vs. \quad H_1 : \alpha_1 < 0.$$

Em particular, continuando a assumir a exogeneidade fraca do(s) regressor(es), o teste pode ser efectuado no contexto dos modelos ADL, como no modelo ADL(1,1), por exemplo,

$$y_t = \mu + \alpha y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \epsilon_t,$$

escrito sob a forma de Bardsen,

$$\Delta y_t = \mu + (\alpha - 1)y_{t-1} + \beta_0 \Delta x_t + (\beta_0 + \beta_1)x_{t-1} + \epsilon_t,$$

trata-se de testar a condição necessária (mas não suficiente) de estabilidade

$$\begin{array}{ll} H_0 : \alpha = 1 \Leftrightarrow & -A(1) = 0, \quad vs. \\ H_1 : \alpha < 1 \Leftrightarrow & -A(1) < 0. \end{array}$$

E de forma semelhante à dos segundos termos para modelos ADL com mais regressores[10] e de ordens ADL mais elevadas.

Note-se, no entanto que, sob H_0, o coeficiente em teste é um de uma variável $I(1)$; mesmo escrito sob a forma de MCE, o coeficiente não pode escrever-se como coeficiente de uma variável $I(0)$ de média nula, pelo que o importante resultado de Sims, Stock e Watson não pode ser empregue. Isto é, a distribuição da estatística de teste não é *standard*.

Após algumas tentativas com pouco sucesso para tabelar a distribuição da estatística, só com o artigo de Ericsson e MacKinnon (2002) foram publicadas tabelas bastante completas de valores críticos. Essas tabelas podem ser consultadas para cada caso. Todavia, um valor crítico grosseiro, aproximado, para testes com dimensão de 5% é dado por

$$VC(0.05) \approx -3.0 - 0.2m - 0.3(nd - 1),$$

[10] Nestes casos tem ainda que verificar-se a hipótese de não existir cointegração entre os regressores.

onde m representa o numero total de variáveis $I(1)$ envolvidas $(g+1)$ e nd representa o número de termos determinísticos.

Com alguma frequência, os testes de EG não darão evidência de cointegração mas o teste t-MCE dará. Poderia ser um caso de distorções de dimensão do segundo teste mas, em geral, não é; trata-se, antes, de um problema de escassez de potência dos testes de EG. A que se deve esta potência reduzida? Deve-se ao que se pode considerar como restrições de factores comuns, a restrições sobre a dinâmica das interacções entre as variáveis que, caso não sejam válidas, retiram potência aos testes.

Sob o ponto de vista da modelação ADL a situação é bastante clara. A título de exemplo, considere-se o modelo ADL(1,1), sem constante para simplificar, escrito sob a forma MCE:

$$\Delta y_t = (\alpha - 1)(y_{t-1} - \lambda x_{t-1}) + \beta_0 \Delta x_t + \epsilon_t,$$

onde $\lambda = (\beta_0 + \beta_1)/(1 - \alpha)$. Subtraindo $\lambda \Delta x_t$ a ambos os membros, obtém-se

$$\underbrace{\Delta(y_t - \lambda x_t)}_{\Delta u_t} = \underbrace{(\alpha - 1)}_{\phi} \underbrace{(y_{t-1} - \lambda x_{t-1})}_{u_{t-1}} + \underbrace{(\beta_0 - \lambda)\Delta x_t + \epsilon_t}_{v_t},$$

onde, note-se que só se pode assegurar que v_t é um processo ruído branco se $\beta_0 = \lambda$, isto é, se os multiplicadores de curto e de longo prazo forem idênticos, o que constitui uma restrição dinâmica muito forte, típica dos modelos estáticos com erros autocorrelacionados; em geral, esta restrição não deverá verificar-se, pelo que a estimação da equação do teste de EG deverá ser ineficiente e, portanto, deverá dar origem a um teste pouco potente.

De regresso ao exemplo que nos tem acompanhado, note-se que se o modelo ADL incluir apenas as variáveis em nível LCP, LRD e INF, mas não LSR pois optou-se, neste caso, por incluir apenas a dummy de impulso $A75$, obtém-se uma estatística $t_{MCE} = -4.26$. Como

$$VC(0.05) \approx -3 - 0.2 \times 3 - 0.3(1 - 1) = -3.6,$$ [11]

obtém-se forte evidência contra a hipótese nula, isto é, forte evidência de cointegração, contrariando a inferência proporcionada pelo teste de Engle-Granger. Pela

[11] O valor crítico a 5% "exacto" (estimado) obtidos das tabelas de Ericsson e MacKinnon (2002) é cerca de -3.64.

razão apontada, deverá ser a primeira destas inferências a prevalecer. De resto, este é um conflito que se observa com alguma frequência, com o teste de EG a revelar-se pouco potente.

Refira-se, por outro lado, que se incluir ainda no ADL a variável em nível LSR (desfasada um período), a evidência de cointegração subsiste mas de forma mais fraca pois a estatística t_{MCE} reduz-se para -4.12 (e o novo valor crítico aumenta em valor absoluto).

7 Estimação do MCE Condicional

A estimação do MCE condicional é o único aspecto que ainda resta para concluir o estudo restrito que se tem vindo a fazer.

7.1 Estimação num só passo

Por si só, o nome deste método parece estranho. O mistério é desvendado pelo facto de o método a abordar alternativamente a este ser o método "em dois passos".

Não existe aqui qualquer novidade. Trata-se de estimar o modelo ADL, de preferência sob a forma de Bardsen, com a estimação dos parâmetros de cointegração (do longo prazo) e da dinâmica de curto prazo a ser efectuada simultaneamente.

Recuperando o exemplo abordado no contexto do estudo dos modelos ADL mas dando preferência à versão que inclui a *dummy* de impulso $A75$ e que repele a variável LSR [12], tem-se

$$\begin{aligned}\widehat{\Delta LC}_t = \;& 0.061 + 0.062A75 - 0.271LCP_{t-1} + 0.258LRD_{t-1} - 0.003INF_{t-1} \\ & +0.403\Delta LC_{t-1} + 0.508\Delta LR_t - 0.267\Delta LR_{t-2}.\end{aligned}$$

7.2 Método dos 2 passos

O método dos 2 passos foi o método de estimação originalmente proposto por Engle e Granger e dominou as aplicações empíricas durante alguns anos.

No primeiro passo é estimado (com o OLS) o vector de cointegração com base na regressão estática; mesmo que se justifique a inclusão de um termo de tendência determinística, este pode ser omitido para tentar reforçar as propriedades de potência

[12] Desta forma, também a estimativa do vector de cointegração difere da apresentada no quadro 2.

do teste de EG. O objectivo é o de obtenção dos resíduos OLS, e_t ou $\widehat{u}_t$, estimativas dos erros de equilíbrio.

No segundo passo, essas estimativas são inseridas na equação do MCE, como se se tratassem dos verdadeiros valores desses erros:

$$\Delta y_{1,t} = \delta + \phi e_{t-1} + \sum \theta_i \Delta y_{1,t-i} + \sum \boldsymbol{\gamma}_j' \mathbf{y}_{2,t-j} + \epsilon_{1t}.$$

Dada a propriedade de superconsistência do OLS no primeiro passo, a substituição dos parâmetros de cointegração pelos seus estimadores não afecta, assintoticamente, as propriedades de estimação e a validade dos métodos de inferência. Todavia, note-se que estudos de simulação mostraram que o estimador num só passo parece ter melhores propriedades de não enviesamento e de eficiência que este método.

O teste $t - MCE$ também se poderá basear neste modelo. Todavia, dado o enfraquecimento da eficiência em pequenas amostras, também a potência tenderá a ser baixa.

Note-se ainda que não é obrigatório que as estimativas dos erros de equilíbrio sejam obtidas, no primeiro passo, pelo OLS sobre a regressão estática. Qualquer outro estimador do vector de cointegração pode ser empregue no primeiro passo: por exemplo, o estimador DOLS ou, até, o estimador de Johansen, embora neste contexto este seja contra-intuitivo, fazendo pouco sentido.

8 Leituras Adicionais

- A nível básico/introdutório: Hendry e Juselius (2000) contém uma abordagem simples e pedagógica, de divulgação; Maddala e Kim (1998) contém referências a aspectos mais avançados mas sempre com uma abordagem pouco profunda; também é útil para abrir outras pistas de leitura.

- A um nível mais avançado: Monteiro e Lopes (2010), pelo exemplo de aplicação e contendo uma componente importante de análise multi-equacional, que pode ser saltada.

  O livro de Juselius (2006) é claramente aconselhável a quem pretender iniciar-se na análise multi-equacional e não é propriamente muito avançado ou complexo; todavia, é bastante extenso.

  Também não pode deixar de se recomendar a leitura das páginas 571 a 582 de Hamilton (1994), pela bela apresentação que desagua no Teorema de Representação de Granger, requerendo a abordagem vectorial, a empreender em MA II.

9 Exercícios

1. Indique, justificando a sua resposta, se a afirmação que se segue é verdadeira ou falsa. Suponha que as séries y_{1t}, y_{2t} e y_{3t} são (todas) I(1); então, pode não existir nenhum vector $\beta' = [\beta_1 \;\; \beta_2]'$ tal que

$$\beta_1 y_{1t} + \beta_2 y_{2t} = u_t \sim I(0),$$

mas pode existir um vector $\alpha' = [\alpha_1 \;\; \alpha_2 \;\; \alpha_3]'$ tal que

$$\alpha_1 y_{1t} + \alpha_2 y_{2t} + \alpha_3 y_{3t} = w_t \sim I(0).$$

2. Considere que $[y_{1t} \;\; y_{2t} \;\; y_{3t}]' \sim I(1)$. Mostre que a característica de cointegração deste vector não pode ser igual a 3.

3. Considere que o processo $\{(y_t, x_t)\}$ é gerado pelo sistema

$$\left\{ \begin{array}{rcl} \alpha y_t - x_t & = & u_t, \;\; u_t = u_{t-1} + \epsilon_{1t}, \\ y_t - \gamma x_t & = & v_t, \;\; v_t = 0.4 v_{t-1} + \epsilon_{2t}, \end{array} \right.$$

com $\alpha\,\gamma \neq 1$ e $\epsilon_{jt} \sim iid(0, \sigma_j^2)$, $j = 1, 2$. Mostre que $(y_t, x_t) \sim CI(1, 1)$.

4. Segundo a teoria ou hipótese do rendimento permanente ("*permanent income hypothesis*"), as variáveis consumo e rendimento disponível das famílias devem ser cointegradas. Empregando os dados trimestrais do ficheiro `PIH_US.xls` para a economia dos E.U.A. e para o período 1984:1–2005:4, analise a aderência da hipótese à realidade empregando uma (ou mais) metodologia(s) que considere apropriada(s).

5. (De um exame de 2018.) Fez-se um estudo de simulação com o PGD dado por

$$x_t = x_{t-1} + \epsilon_{1t}, t = 1, 2, \ldots, 100,$$

$$y_t = \left\{ \begin{array}{ll} x_t + \epsilon_{2t}, & t = 1, 2, \ldots, 50 \\ (1 - \frac{\delta - 2}{10}) x_t + \epsilon_{2t}, & t = 51, 52, \ldots, 100, \; \delta = 2, 4, \ldots, 12. \end{array} \right.$$

com

$$\begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix} \sim iid\mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0.3 \end{bmatrix} \right).$$

Com as 100 observações, estimaram-se as equações $y_t = \beta_1 + \beta_2 x_t + u_t$ e, com os resíduos OLS dessas regressões (e_t), também as equações $\Delta e_t = \theta e_{t-1} + v_t$, para efectuar o teste de $H_0 : \theta = 0$ vs. $H_1 : \theta < 0$, empregando RC $= \{t_\theta : t_\theta < -3.37\}$. No quadro seguinte apresentam-se as percentagens de rejeições desse teste em 10000 réplicas:

δ	2	4	6	8	10	12
% rej.	100.0	90.8	63.1	44.4	30.9	23.3

a) Explique o PGD e o procedimento estatístico adoptado, b) bem como o(s) objectivo(s) do estudo e c) retire a(s) conclusão(ões) apropriada(s).

6. (De um exame antigo.) Suponha que, no início de um trabalho empírico, foram obtidos resultados que indicam que $y_t \sim I(1)$. Suponha ainda que, no final desse trabalho, foi apresentado como produto final, o seguinte modelo estimado:

$$\begin{array}{llllll} \widehat{\Delta y_t} = & 0.003 & -0.345 y_{t-1} & -0.332 x_{t-1} & +0.010 z_{t-1} & +0.653 \Delta x_t \\ & (0.001) & (0.068) & (0.102) & (0.031) & (0.212) \\ & +0.091 \Delta x_{t-1} & -0.232 \Delta z_{t-1} & +0.125 \Delta y_{t-1} & & \\ & (0.035) & (0.089) & (0.052) & & \end{array}$$

(a) Que conclusões intermédias terão sido obtidas? (b) Escreva e interprete o modelo sob a forma de MCE, não se esquecendo de indicar a (eventual) relação de equilíbrio de longo prazo estimada. (c) Teste a existência dessa relação e indique as hipóteses que necessita assumir para que esse teste seja válido (assintoticamente).

7. Considere o modelo ADL(1,1)

$$y_t = \mu + \alpha y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + \epsilon_t,$$

e suponha que tanto y_t como x_t são $I(1)$, sem deriva, e cointegradas ($(y_t, x_t) \sim CI(1,1)$). É possível fazer inferência da forma habitual, com estatísticas-t válidas assintoticamente, usando a aproximação pela normal, sobre os coeficientes

a) β_0? b) β_1? c) α?

d) e sobre os 3 coeficientes conjuntamente (com uma estatística-F)?

8. Reconsidere o exercício 5 do cap. 3. Refaça-o considerando:

   – que todas as séries envolvidas são I(1);

   – que um dos principais objectivos é obter provas estatísticas da presença de cointegração e escrever (e interpretar) o modelo final como um MCE;

   – que outro objectivo é o de obter previsões de boa qualidade a um passo;

   – que o modelo final deverá ser considerado adequado de acordo com vários testes de especificação (isto é, "congruente com os dados", na terminologia de David F. Hendry).

10 Bibliografia

Baneerjee, A., Dolado, J., Galbraith, J. W. e Hendry, D. F. (1993), *Co-Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data*, Oxford University Press.

Boswijk, H. P. (1994), *Unit Roots and Cointegration: Statistical Analysis and Asymptotic Theory*, Tinbergen Institute and University of Amsterdam.

Davidson, R. e MacKinnon, J. G. (1993), *Estimation and Inference in Econometrics*, Oxford University Press.

Engle, R. F. e Granger, C. W. J. (1987), Cointegration and error correction: representation, estimation and testing, *Econometrica*, 55, 251-76.

Engle, R. F. e C. W. J. Granger (eds.) (1991), *Long-run Economic Relationships*, Oxford University Press.

Ericsson, N. R. (1994). Testing exogeneity: an introduction, *in* Ericsson, N. R. e Irons, J. S. (eds.) (1994), *Testing exogeneity*, Oxford University Press, pp. 3-38.

Ericsson, N. R. e MacKinnon, J. G. (2002), Distributions of error correction tests for cointegration, *Econometrics Journal*, 5, 285-318.

Granger, C. W. J. (1986), Developments in the study of cointegrated economic variables, *Oxford Bulletin of Economics and Statistics*, 48, 213-28.

Granger, C. W. J. (1991), Some recent generalizations of cointegration and the analysis of long-run relationships, *in* Engle, R. F. e C. W. J. Granger (eds.) (1991), *Long-run Economic Relationships*, 277-287.

Hamilton, J. D. (1994), *Time Series Analysis*, Princeton University Press.

Hayashi, F. (2000), *Econometrics*, Princeton University Press.

Hendry, D. F. e Juselius, K. (2000), Explaining cointegration: Part I, *The Energy Journal*, vol. 0(1), 1-42[13].

[13] Tratando do caso multi-equacional, a parte II deste artigo foi publicada em 2001 na mesma revista.

Johnston, J. e DiNardo, J. (1997), *Econometric Methods*, McGraw-Hill.

Juselius, K. (2006), *The Cointegrated VAR Model*, Oxford University Press.

Lütkepohl, H. (2004), Vector autoregressive and vector error correction models. In: Lütkepohl, H. and Krätzig, M. (eds.), *Applied Time Series Econometrics*, Cambridge University Press, 86-158.

Maddala, G. S. e Kim, I.-M. (1998), *Unit Roots, Cointegration and Structural Change*, Cambridge University Press.

Monteiro, O. S. M. e Lopes, A. S. (2010), Short- and long-run tests of the expectations hypothesis: the Portuguese case, *Applied Economics Quarterly*, 56/3, 257-80.

Murray, P. M. (1994), A drunk and her dog: an illustration of cointegration and error correction, *The American Statistician*, 48(1), 37-39.

Ng, S. e Perron, P. (1997), Estimation and inference in nearly unbalanced nearly cointegrated systems, *Journal of Econometrics* 79 (1), 53-81.

Phillips, P. C. B. e Ouliaris, S. (1990), Asymptotic properties of residual based tests for cointegration, *Econometrica*, 87 (1), 165-93.

Stewart, J. e Gil, L. (1998), *Econometrics*, 2nd ed., Prentice Hall.

Stock, J. H. (1987), Asymptotic properties of least squares estimators of cointegrating vectors, *Econometrica*, 55, 1035-56.

Watson, M. (2010). Cointegration, *in* Durlauf, S. N. e Blume, L. E. (eds), *Macroeconometrics and Time Series Analysis*, pp. 53-59.